DESIDOC Journal of Library & Information Technology, Vol. 44, No. 5, September 2024, pp.277-283, DOI : 10.14429/djlit.44.06.19891

Implementing Free Persistent Identifiers in a Scientific Journal Management System Alternatively to DOI

Yasiel Pérez Vera* and Álvaro Fernández Del Carpio

Software Engineering Department/Universidad La Salle, Arequipa - 04011, Perú
**E-mail: yperez@ulasalle.edu.pe*

ABSTRACT

Scientific research is the tool that drives the development of society in all areas, and research articles are the primary means of disseminating advances in science and technology. The main way to identify a scientific article is through the persistent identifiers that journals assign to articles. There are a variety of persistent identifiers, with the digital object identifier being among the most widely used. The use and assignment of most persistent identifiers have an associated financial cost. Journals sometimes have difficulty covering this cost or transferring it to the authors of the articles, causing barriers to publishing scientific articles. This work aims to implement free persistent identifiers in a scientific journal management system. To accomplish this, software components for managing the persistent identifiers archival resource key and persistent uniform resource locator are implemented for the open journal system. The results of implementing these persistent identifiers in a scientific journal are analysed, and the implementation by the scientific community of 14 countries in 32 journals. Implementing free persistent identifiers in scientific journals allows for greater visibility and savings in publication costs.

Keywords: Persistent identifier; Open journal system; Archival resource key; Persistent uniform resource locator

1. INTRODUCTION

Web-based persistent identifiers have been in existence for over two decades. These identifiers were developed to address challenges arising from the decentralised nature of the internet, which led to difficulties in maintaining a persistent record of science and causing issues such as "link rot." Many journals, scientific societies, and funding agencies require open access to data in articles, and some publishers have introduced specialised data-focused journals to emphasise their long-term utility, ensuring stability, completeness, permanence, and high quality[1]. The introduction of persistent identifiers (PIDs) aimed to separate the identity of an object from its location on the web, ensuring global uniqueness and unambiguous identification of resources[2]. PIDs have traditionally been a subject of focus in disciplines like library science and scientific publishing[3]. Various PID systems, such as the Handle System, Persistent URL (PURL), and Archival Resource Key (ARK), have been established to offer enduring references to digital objects. These systems are widely utilised and commonly supported by digital archival and research information systems[4].

An example of the current use of web-based persistent identifiers is in digital libraries and repositories such as arXiv, a preprint repository for scientific papers in physics, mathematics, and computer science. arXiv uses ARK identifiers to ensure the long-term accessibility and reliability of the documents it hosts. Another example is the California Digital Library (CDL), which uses ARK identifiers to manage and preserve digital collections. These identifiers provide stable links to resources, ensuring they remain accessible even if the underlying web address changes[5].

PIDs address issues related to accessibility, integration, and interoperability in the scholarly landscape. By providing a unique and persistent link to a determined resource, PID allows resources to be located, visible and cited accurately in publications. They facilitate easier tracking and measuring of scholarly work, enabling researchers to receive credit for their data and assert their intellectual property[6-7]. In general, using PIDs enhances transparency, reproducibility, and accountability within scientific research[7] and facilitates the sharing and reusing of research data, leading to faster solutions to critical problems[7,9]. Overall, PIDs are essential for promoting open science and improving metadata quality through linking connected metadata sources.

For new scientific journals with limited financial resources, the expense of paid persistent identifiers may present a significant barrier to adopting these services. The cost associated with paid identifiers could force new journals to prioritise essential operational expenses over investing in persistent identifier services. Moreover, the financial burden of paid persistent identifiers may hinder new journals from fully embracing the best research citation and data-sharing practices.

Received : 07 February 2024, Revised : 10 July 2024
Accepted : 16 July 2024, Online published : 24 September 2024

Scientific journals often pass the cost of PIDs onto authors to be paid indirectly as part of the publication process. This practice can burden authors financially, particularly those from low-income or emerging countries without institutional funding. As a result, it may discourage authors from publishing their work, thereby limiting the dissemination of knowledge.

This work aims to implement free persistent identifiers in a scientific journal management system. Unlike the paid DOI system, the plugins that will be implemented offer a cost-effective alternative, as ARK and PURL identifiers are free. The plugins stand out because they seamlessly integrate with the platform Open Journal System (OJS), a widely used journal management platform, addressing a significant gap in the market where existing solutions are either paid or lack compatibility with OJS. This advancement promotes financial efficiency and enhances the accessibility and sustainability of digital scholarly communication.

This paper is structured as follows: In Section 2, we presented the literature review; in Section 3, the materials and methods used for this research are described in detail. Then, in Section 4, the results are presented. In Section 5, we presented the discussions. Finally, the conclusions are established in Section 6.

2. LITERATURE REVIEW

This section presents research papers extracted from the literature about applying PIDs; the most relevant ones are described as follows.

The significance of PIDs in establishing a trusted ecosystem for research data was discussed in[7]. The authors explore the technical and human infrastructure needed to achieve interoperability across the PID landscape, emphasising the advantages of scalable identifier systems in capturing research activity. Additionally, they offer insights into how interconnectivity was facilitated technically within the research record and address the social and human challenges associated with increasing adoption.

The adoption of PIDs in scholarly publications focused on datasets, software, and research facilities are described by[7]. Through various case studies, the research evaluates reference resources assigned to persistent identifiers, the evolution of citations, and the advocacy of referencing behavior. They noted the importance of consistent referencing practices, the role of journal editorial policies, and the positive trends in research infrastructure citation. The findings underscored the data citation principles and tools, emphasising the need for continued education and overcoming efforts to promote effective referencing of research infrastructures.

The recent increased use of PIDs for physical aspects enhances the Findability, Accessibility, Interoperability, and Reusability (FAIR principles)[8]. The role of institutional repositories is crucial in managing physical samples and artifacts, ensuring their preservation and access. The author suggests that funders, research communities, and professional organisations should establish leading practices and data standards to promote the adoption of PIDs for physical aspects and how to support them.

In[10], the importance of PIDs as foundational elements in the research information infrastructure is discussed. Adopting PIDs can lead to improved access to information, collaboration opportunities, reduced administrative overhead, and increased trust in scholarship and research. The requirement for PID metadata to adhere to FAIR principles and to transparent sourcing of information is also presented. The benefits of expanded PID adoption, such as improved recognition for researchers' contributions and enhanced capabilities for research organisations and the wider community, were highlighted as essential aspects to be considered. Efforts to enhance the adoption of PIDs are gaining attention at the national level, as demonstrated by initiatives like the French national plan for open science.

3. METHODOLOGY

This section will describe the main characteristics and concepts associated with persistent identifiers. The most commonly persistent identifiers will be listed and detailed, including the free PIDs ARK and PURL. Additionally, the main features and functionalities of the OJS software will be outlined.

The methodology of this work involves several key steps to ensure the successful implementation and integration of the developed plugins for the OJS platform. Initially, the focus is on developing and testing the plugins that enable ARK and PURL identifiers. Following the development, these plugins are integrated into the OJS repositories, adhering to the established code review processes of the OJS community. This step involves community members thoroughly examining and validating the code to ensure security, functionality, and compatibility. Once the community approves the plugins, they are integrated into the OJS plugin gallery, making them available for widespread use.

The next phase involves installing the approved plugins in a scientific journal, which will serve as a case study to evaluate the impact of the developed plugins. This case study is essential to assessing the practical benefits and cost savings of using ARK and PURL identifiers instead of DOIs. The economic impact will be evaluated by comparing the costs before and after implementing these plugins, highlighting the financial efficiency achieved through their use.

Additionally, the case study will measure the journal's visibility by analysing various indicators, such as the number of users, the number of sessions with interaction, and the geographical distribution of visitors to the journal's website. Other metrics to be evaluated include the number of published articles, the number of articles with international participation, and the number of citations received by the journal's articles. These metrics will provide a comprehensive understanding of the plugins' impact on the journal's reach, engagement, and academic influence, thereby validating the effectiveness of the implemented solution.

3.1 Persistent Identifiers

Persistent identifiers are mechanisms for the unambiguous, persistent, and functional identification of digital objects resulting from scientific research and communication practices[11]. They enable formal and unambiguous referencing of documents, including journal articles, books, book chapters, conference proceedings, communications, software, and videos. They allow these digital objects to be identified, represented, and used, facilitating the reuse, citation, and socialisation of scientific and academic productions[12].

According to the ODIN Consortium[13], a PID must have the following characteristics: name, globally unique, persistent over time, resolvable as a URI, managed by an organisation, with metadata to describe relevant properties, interlinkable and interoperable with other identifiers.

On the other hand, the entities that must be included for PIDs are agents, resources, rights statements, events, and projects.

Several persistent identifiers are used to uniquely and permanently identify and reference digital resources[14]. Here are some examples:

- DOI (Digital Object Identifier) is a unique PID assigned to digital documents such as journal articles, technical reports, e-books, research data, and other online resources. It is the PID most used by the scientific community.
- Handle is a resource identification and resolution system that provides unique identifiers for digital objects such as documents, images, and other resources.
- ARK is a unique PID for digital resources in digital preservation and collection management.
- PURL is a unique PID for digital resources that uses a web address to provide permanent access to a resource.

Although financial aspects are an important consideration when selecting an identifier, some PIDs like ARK and PURL, for example, provide some key benefits over DOI:

- ARK can include internal information within its structure, enhancing its informativeness[15].
- The ARK scheme is more lightweight and less packaged than DOI regarding the participation model[16].
- Concerning other PIDs, the ARK qualifier is very useful for identifying components of resource variants. Differently, DOI does the identification task in a manner less controlled[17].
- Unlike URLs, PURLs, and URNs, the ARK offers three components: the object itself, metadata, and a provider's commitment[18].
- An advantage of ARKs over DOIs is the ability to retrieve unresolved metadata5 directly.
- Unlike ARK, DOI, Handle, and PURL necessitate that their respective centralised system provide resolution and other services. Also, DOI and Handle have metadata requirement[15].

3.2 Archival Resource Key

An ARK is a multipurpose PID for information objects of any type. It is expressed in the form of a link. An ARK contains the ark:/ tag after the hostname or resolver and, finally, the resource identifier. The generic structure of an ARK is https://NMA/ark:/NAAN/ID[19]. Figure 1 shows the ARK's structure.

```
  https://example.org/ark:/12345/x54xz321/s3/f8.05v.tiff
  \________________/ \__/ \___/ \______________________/
           |           |     |              |
           |  ARK Label      |          Object ID
           |                 |
Name Mapping Authority (NMA) |
                             |
           Name Assigning Authority Number (NAAN)
```

Figure 1. Structure of ARK persistent identifier[5].

ARKs are the only conventional, no-deposit, no-payout tokens that you can register to use in about 48 hours. DOIs, Handles, and PURLs require resolution and other services from their respective centralised systems (repositories). ARKs are unusual in being decentralised. More than 600 registered organisations worldwide have created an estimated 3.2 billion ARKs; as with URLs, no one has paid an identification fee to make them[20].

3.3 Persistent Uniform Resource Locator

A PURL is a Uniform Resource Locator (URL) often used to reference a specific resource that changes address over time from the same address. PURLs redirect to web clients and handle the URL resolution process. The PURL concept enables widespread Internet preservation of the Uniform Resource Identifier (URI). PURLs allow third-party control over both URL resolution and resource metadata layout[21].

If a web resource changes location, a PURL can be updated pointing to it. A user of a PURL always uses the same web address, although they may have moved the resource in question. Editors can use PURLs to manage their information space or web users to manage theirs. A PURL service is independent of who publishes the information. PURL services thus enable the management of hyperlink integrity[22].

3.4 Open Journal System

OJS is an editorial management system that can be implemented to manage online magazines. It was designed to reduce the time and energy spent on administrative tasks associated with journal publishing and improve record-keeping and editorial processes' efficiency[23]. It aims to strengthen the academic quality of journal publishing through a series of innovations, including enhancing the reader experience, making journal policies transparent, and improving indexing[24].

OJS enables the complete editorial process, from article submission by authors to the assignment of responsible editors, the peer review process, style correction, layout, and the publication with the curation of corresponding metadata[25].

3.5 Implementation of ARK and PURL Persistent Identifiers in OJS

Plugins must be developed to implement the ARK and PURL PIDs into OJS. Plugins are software components that can be added to OJS to expand its functionality. To develop a plugin for OJS, developers must adhere to the guidelines and rules established by the OJS community. The built-in OJS plugin gallery provides access to externally created plugins that may not come pre-installed with your OJS setup but can be downloaded and activated.

Once a software developer implements plugins for ARK and PURL identifiers, they must be reviewed by the OJS community. This process usually takes several weeks from when the source code is uploaded to GitHub and the request to review the plugins is made. Once the code is bug-free, the plugins can be published to the OJS community through the plugin gallery.

There are several categories of plugins in OJS, and each of these categories has different functions. The Public Identifier (PubId) category was used to develop the ARK and PURL identifier plugins. These plugins have features associated with assigning persistent identifiers to objects such as articles, issues, and galleys. Figure 2 shows the classes' inheritance related to the PubId type plugins developed in OJS.

In addition to the developed PubId-type plugins, a Gateway-type plugin was implemented. This Gateway-type plugin allows you to act as a resolver within the OJS backend. Given a PID of any type, that plugin can resolve the web address of the provided persistent identifier (DOI, ARK, PURL, URN). Together with the previous two plugins, this component creates an ecosystem of software components within OJS for implementing PIDs.

The OJS platform is highly secure due to its robust community-driven approach to software development and maintenance. The open-source nature of OJS allows a broad community of developers and researchers to continuously review and audit the source code, ensuring that any vulnerabilities are quickly identified and addressed. Plugins developed for OJS, including those enabling ARK and PURL identifiers, undergo rigorous integration testing to ensure compatibility and security within the platform. Additionally, measures are in place to protect against backdoors, including thorough code reviews, automated testing, and adherence to best practices in secure software development, thereby providing a secure and reliable environment for scholarly publishing.

Also, the plugins developed guarantee availability and interoperability through several key features. These plugins ensure that digital content remains persistently accessible, as ARK and PURL identifiers are designed to provide stable, long-lasting links that are not affected by changes in web addresses. Interoperability is achieved through adherence to widely accepted standards and protocols, allowing seamless integration with various digital libraries, repositories, and other scholarly communication systems. Furthermore, these plugins are regularly updated and tested within the OJS community, ensuring they remain compatible with evolving technological environments and maintain consistent performance across different systems and platforms.

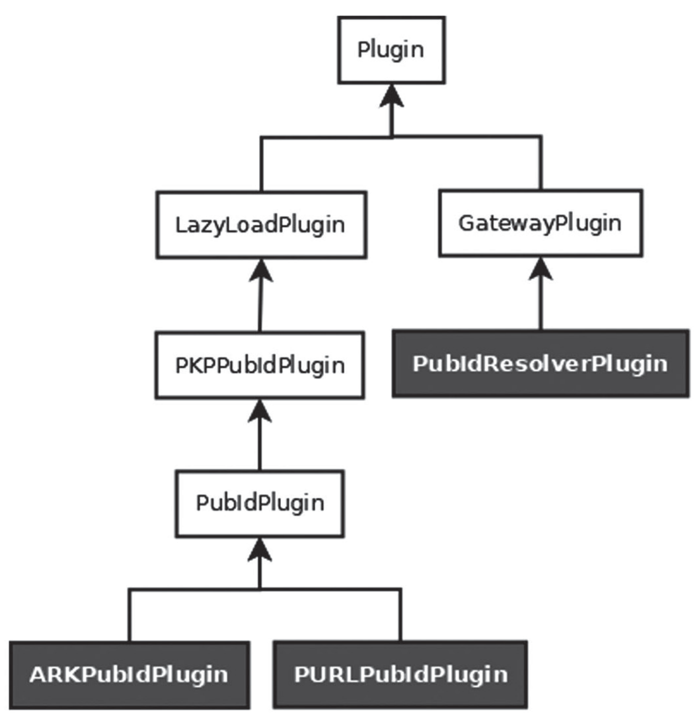

Figure 2. Inheritance of classes associated with ARK and PURL plugin.

4. RESULTS

The free PIDs ARK and PURL were integrated into the OJS platform of the scientific journal "Innovation and Software" of La Salle University. Each issue, article, and galley were assigned an ARK and PURL identifier. A NAAN prefix was requested from the ARK Alliance for ARK identifiers, which assigned the prefix ark:/42411 to the journal "Innovation and Software". Once the prefix was obtained, the plugins developed for the OJS platform were configured to assign the ARK persistent identifiers to each object following a pattern established by the publisher. The free N2T service (https://n2t.net), provided by the ARK Alliance, was used as a global resolver. The plugin configuration for assigning and managing ARK identifiers is displayed in Fig. 3.

The process for configuring the PURL PID plugin was similar to that of the ARK plugin. In this case, the URLs of the PURL identifier had to be beforehand registered in the Internet Archive's PURL service. The resolver was configured with the services provided by the Internet Archive, using the URL https://purl.org.

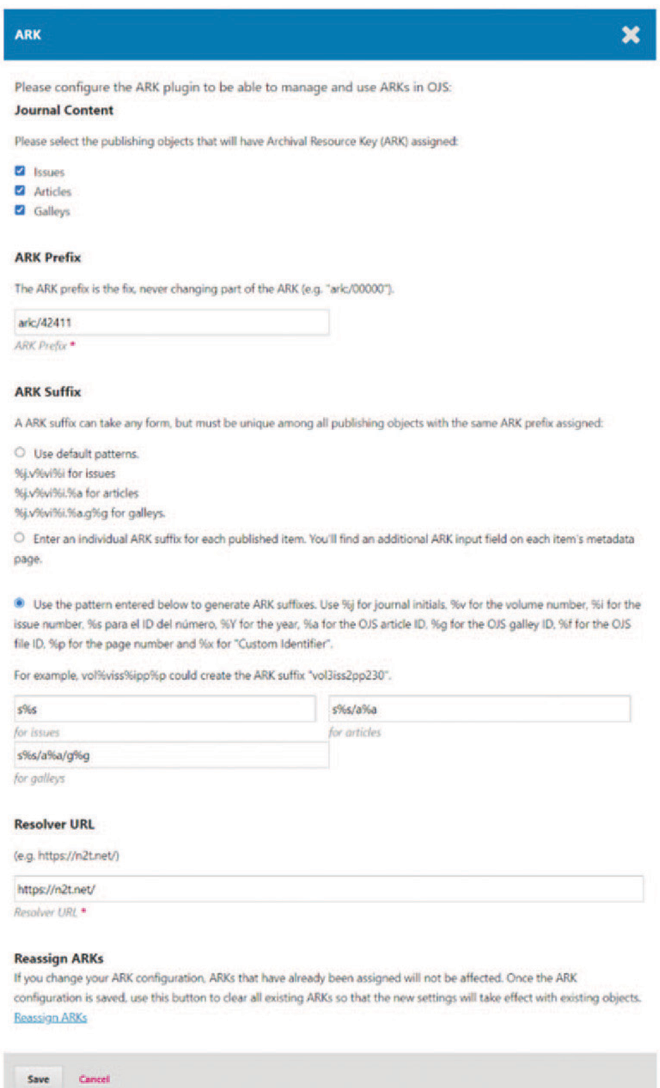

Figure 3. Configuration of ARK plugin in OJS.

Figure 4 depicts the landing page of an article where an ARK and a PURL are assigned.

Because free PIDs were implemented in the scientific journal "Innovation and Software" of La Salle University in 2023, the visibility indicators for 2023 are compared with previous years with data available. The 2024 data presented is from January to June. The website considerably increased user visits, sessions, and user interactions. Additionally, there was an uptick in the number of articles submitted and international author participation. Finally, the number of citations reported by Google Scholar also increased significantly. Table 1 shows a comparison between visibility data from 2021 to 2024.

5. DISCUSSION

From an editorial management point of view, there were no costs associated with PIDs. The average cost of PIDs has two elements: membership to the agency that manages the DOIs and the cost per DOI generated.

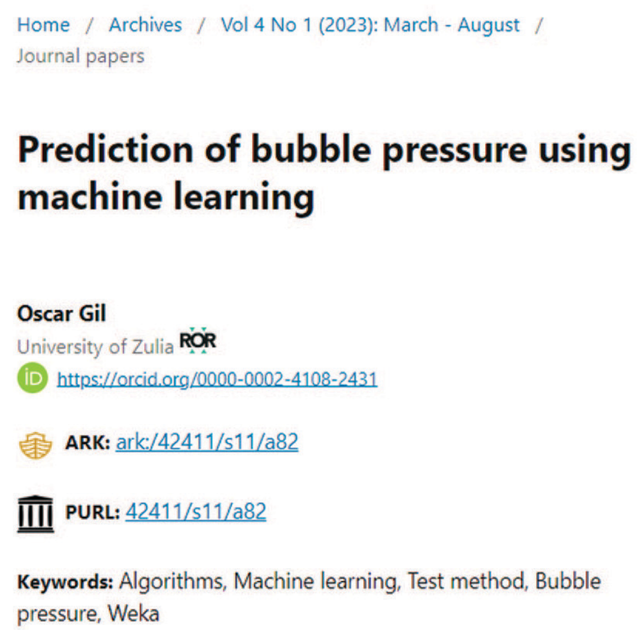

Figure 4. An example of an article with ARK and PURL identifiers.

Table 1. Comparison of visibility data

Elements	2001	2022	2023	2024
Users	113	4,892	9,324	7,186
Sessions	175	7,996	15,288	11,413
Sessions with interaction	108	4,988	9,153	6,979
Visitor's countries	20	71	81	78
Articles published	13	19	29	12
Articles received	18	25	38	17
Articles with international authors	3	3	4	2
Google scholar citations	35	33	49	29

Membership fees typically range around USD 275, while the cost per DOI assigned is usually USD 1. With the implementation of free persistent identifiers, La Salle University estimates annual savings of USD 350.

According to the ARK Foundation's Registry, 41 scientific journals have registered and applied for an ARK prefix. Of these, 32 scientific journals have carried out the implementation proposed by this work correctly. Journals are working correctly with the proposed implementation of the following countries: Peru (2), Brazil (9), Cuba, Algeria, Argentina (2), Mexico, Morocco, Turkey, Ecuador (4), Kenya, India (5), Costa Rica (2), United Kingdom and Tunisia. This indicates widespread adoption of the proposed implementation of free PIDs for scientific journal management systems across 14 countries.

6. CONCLUSIONS

The present work describes the implementation process of PIDs ARK and PURL in the scientific journal management system of La Salle University. Once persistent identifiers were implemented in the Scientific Journal Management System of La Salle University, it was observed that the production of scientific articles increased considerably. Additionally, the scientific journal "Innovation and Software" of La Salle University achieved greater international visibility. Free PIDs offer an excellent alternative for scientific journals to consider, especially during their initial years of operation.

REFERENCES

1. Cook, R.B.; Vannan, S.K.S.; McMurry, B.F.; Wright, D.M.; Wei, Y.; Boyer, A.G. & Kidder, J.K. Implementation of data citations and persistent identifiers. ORNL DAAC. *Ecological Informatics*, 2016, **33**, 10–6. doi: 10.1016/j.ecoinf.2016.03.003
2. Klump, J. & Huber, R. 20 Years of persistent identifiers - which systems are here to stay? *Data Sci. J.*, 2017, **16**, 1–7. doi: 10.5334/dsj-2017-009
3. Weigel, T.; Kindermann, S. & Lautenschlager, M. Actionable persistent identifier collections. *Data Sci. J.*, 2014, **12**, 191–206. doi: 10.2481/dsj.12-058
4. Sicilia, M.A.; García-Barriocanal, E.; Sánchez-Alonso, S. & Cuadrado, J.J. Decentralised persistent identifiers: A basic model for immutable handlers. *Procedia Computer Sci.*, 2019, **146,** 123–30. doi: 10.1016/j.procs.2019.01.087
5. ARK Alliance. https://arks.org/about/ (accessed on 24 June 2024).
6. Bangert, D. & Frances, M. PIDs to support discovery and citation: Persistent identifier service design and delivery at UNSW Library. *In* 2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL), 19-23 June 2017, Toronto. pp. 5-6. doi: 10.1109/JCDL.2017.7991610
7. Dappert, A.; Farquhar, A.; Kotarski, R. & Hewlett, K. Connecting the persistent identifier ecosystem: Building the technical and human infrastructure for open research. *Data Sci. J.*, 2017, **16**, 1-16. doi: 10.5334/dsj-2017-028
8. Mayernik, M.S. & Maull, K.E. Assessing the uptake of persistent identifiers by research infrastructure users. *PLoS ONE J.,* 2017, **12**(4), 1-15. doi: 10.1371/journal.pone.0175418
9. Plomp, E. Going digital: Persistent identifiers for research samples, resources and instruments. *Data Sci. J.*, 2020, **19**(1), 1-8. doi: 10.5334/dsj-2020-046
10. Meadows, A.; Haak, L.L. & Brown, J. Persistent identifiers: the building blocks of the research information infrastructure. *Insights UKSG J.*, 2019, **32**. doi: 10.1629/uksg.457
11. Freire, N.; Manguinhas, H.; Isaac, A. & Charles, V. Persistent identifier usage by cultural heritage institutions: a study on the europeana.eu dataset. *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2023, 341-348. doi: 10.1007/978-3-031-43849-3_31
12. Gould, M. People, places, and things Persistent identifiers in the scholarly communication landscape. *College and Res. Libr. News*, 2022, **83**(9), 398-402. doi: 10.5860/crln.83.9.398
13. ODIN Consortium, Fenner, M.; Thorisson, G.; Ruiz, S. & Brase, J. (2013). D4. 1 Conceptual model of interoperability. doi: 10.6084/m9.figshare.824314.v1
14. Klein, M. & Balakireva, L. On the persistence of persistent identifiers of the scholarly web. *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2020, 102-115. doi: 10.1007/978-3-030-54956-5_8
15. ARK Alliance. https://arks.org/about/comparing-arks-and-other-identifiers/ (accessed on 24 June 2024).
16. Hilse, H-W. Implementing persistent identifiers. Research and Development Department of the Goettingen State and University Library, 2017.
17. Luo,Y. & Plale, B. Pilot evaluation of Collection API with PID Kernel Information. 2019.
18. Philipson J. Identifying PIDs playing FAIR. *Data Sci.,* 2019, **2**(1–2), 229–44. doi: 10.3233/DS-190024
19. Ananthakrishnan, R.; Chard, K. & D'Arcy, M, *et al.* An open ecosystem for pervasive use of persistent identifiers. *In* ACM International Conference Proceeding Series. 2020, 99-105. doi: 10.1145/3311790.3396660
20. Kunze, J.; Janée, G. & Starr, J. EZID: Easy identifier and metadata management. *In* Proceedings of the International Conference on Dublin Core and Metadata Applications, 1-4 September, 2015, Sao Paulo. pp. 190-191. doi: 10.23106/dcmi.952137156
21. Overton, J.A.; Cuffaro, M. & Mungall, C.J. The OFOCTWG. String of PURLs - frugal migration and maintenance of persistent identifiers. *Data Sci. J.*, 2020, **3**(1), 3-13. doi: 10.3233/DS-190022

22. Duerr, R.E.; Downs, R.R.; Tilmes, C., *et al.* On the utility of identification schemes for digital earth science data: An assessment and recommendations. *Earth Sci. Infor.*, 2011, **4**(3), 139-160. doi: 10.1007/s12145-011-0083-6
23. Ahmadi, D.; Lisnur, W.; Nurrahman, A.A.; Yanuarti, E. & Basudewa, M.I. Improving scientific literacy through management of electronic journal using the "open journal system". *In* AIP Conference Proceedings, 24 August, 2023, Medan. pp. 1-4. doi: 10.1063/5.0158230
24. Koley, M.; Namdeo, S.K.; Suchiradipta, B. & Afifi, N.A. Digital platform for open and equitable sharing of scholarly knowledge in India. *J. Librarianship Infor. Sci.*, 2023, **55**(2), 403-413. doi: 10.1177/09610006221083678
25. Spirin, O.M.; Matviienko, O.V.; Ivanova, S.M.; *et al.* The use of open electronic scientific and educational systems to support the professional activities of research and teaching staff of ukrainian universities and scientific institutions. *In* ACM International Conference Proceeding Series, 18-19 June, 2021, Ruse. pp. 169-176. doi: 10.1145/3526242.3526261

CONTRIBUTORS

Dr Yasiel Pérez Vera is a Lecturer Professor at the School of Software Engineering, La Salle University, Arequipa, Peru. He received a BSc degree in Informatics Science and an MSc in Project Management from the University of Informatics Science, La Habana, Cuba, and a PhD in Business Administration (DBA) from the National University Saint Augustine of Arequipa. He is Editor-in-Chief of the scientific journal *Innovation and Software* at La Salle University and Research Professor and Dean of Postgraduate Studies at La Salle University, Arequipa, Peru.

Dr Alvaro Fernández del Carpio received a PhD degree in Computer Science and Technology from the Carlos III University of Madrid, Spain, and an Official Master in Computer Science and Technology from the Carlos III University of Madrid, Spain. He obtained a BSc degree in Systems Engineering from the Santa María Catholic University of Arequipa, Peru. He completed a Master's in Software Engineering at the Catholic University of Santa María Arequipa, Peru. He is the Head of the Research Department at La Salle University and Research Professor at the La Salle University and Catholic University of Santa María.